IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 26, 2020, accepted January 6, 2021, date of publication January 14, 2021, date of current version January 22, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051633

Social Engineering in Cybersecurity: Effect Mechanisms, Human Vulnerabilities and Attack Methods

ZUOGUANG WANG, HONGSONG ZHU, (Member, IEEE), AND LIMIN SUN
School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China
Beijing Key Laboratory of IoT Information Security Technology, Institute of Information Engineering, CAS, Beijing 100093, China

Corresponding authors: Hongsong Zhu (zhuhongsong@iie.ac.cn) and Zuoguang Wang (wangzuoguang16@mails.ucas.ac.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB0802804, and in part by the Joint Fund of the National Natural Science Foundation of China under Grant U1766215.

ABSTRACT Social engineering attacks have posed a serious security threat to cyberspace. However, there is much we have yet to know regarding what and how lead to the success of social engineering attacks. This paper proposes a conceptual model which provides an integrative and structural perspective to describe how social engineering attacks work. Three core entities (effect mechanism, human vulnerability and attack method) are identified to help the understanding of how social engineering attacks take effect. Then, beyond the familiar scope, we analyze and discuss the effect mechanisms involving 6 aspects *(persuasion, social influence, cognition & attitude & behavior, trust and deception, language & thought & decision, emotion and decision-making)* and the human vulnerabilities involving 6 aspects *(cognition and knowledge, behavior and habit, emotions and feelings, human nature, personality traits, individual characters)*, respectively. Finally, 16 social engineering attack scenarios (including 13 attack methods) are presented to illustrate how these mechanisms, vulnerabilities and attack methods are used to explain the success of social engineering attacks. Besides, this paper offers lots of materials for security awareness training and future empirical research, and the model is also helpful to develop a domain ontology of social engineering in cybersecurity.

INDEX TERMS Social engineering, attack model, working mechanism, human hacking, attack scenario, vulnerability, principle, psychology cognition and behavior, education awareness and training, security.

I. INTRODUCTION

In the context of computer and cyber security, social engineering describes a type of attack in which the attacker exploit human vulnerabilities by means such as influence, persuasion, deception, manipulation and inducing, so as to get classified information, hack computer system and network, obtain unauthorized access to restricted areas, or breach the security goals (such as confidentiality, integrity, availability, controllability and auditability) of cyberspace elements (such as infrastructure, data, resource, user and operation). Succinctly, social engineering is a type of attack wherein the attacker exploit human vulnerability through social interaction to breach cyberspace security [1].

In hacker community, social engineering is a quite popular attack since 1970s. Compared to classical computer attacks such as password cracking by brute-force and software vulnerabilities exploit, social engineering attacks focus the exploitation of human vulnerabilities, to bypass or break through security barriers, without having to combat with firewall or antivirus software by deep coding. In addition, there is not a computer system doesn't rely on humans or involves human factors on earth, and these human factors are obviously vulnerable or can be largely turned into security vulnerabilities by skilled attackers. These inevitable and vulnerable human factors makes social engineering to be a universal cybersecurity threat. For some situations, social engineering attacks may be as simple as making a phone call and impersonating an insider to elicit the classified information. Moreover, with the development of new technology and the formation of new cyber-environment, social engineering threat is increasingly serious. Social Network Sites (SNSs), mobile communication, Industrial Internet and Internet of Things (IoT) generate not only large amounts of sensitive information about people and devices but also more attack channels and a bigger attack surface. Unrestricted office environment (bring your own device, remote office, etc.) leads to the weakening of area-isolation of different security levels

The associate editor coordinating the review of this manuscript and approving it for publication was Xiali Hei.

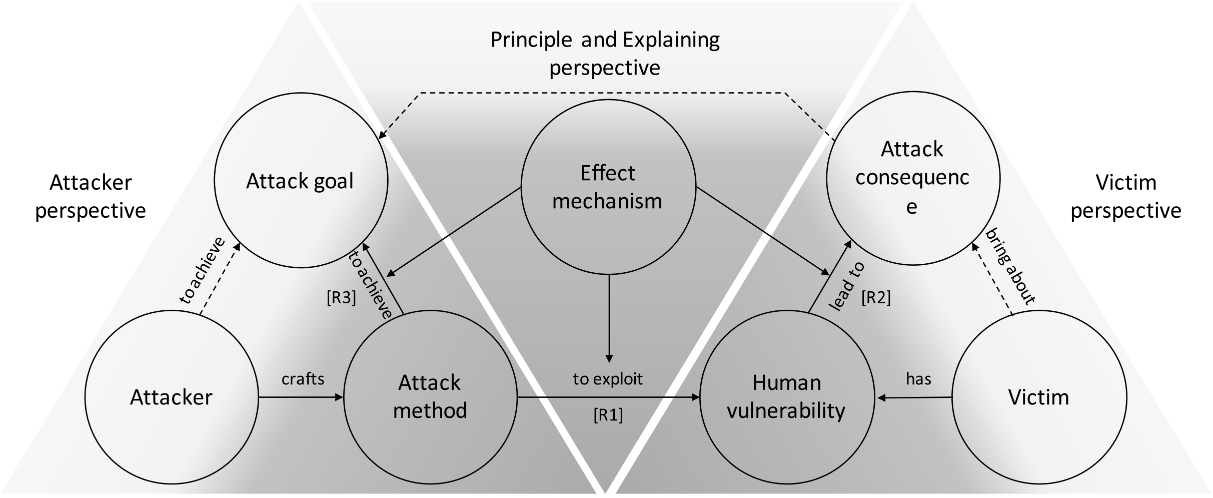

FIGURE 1. A conceptual model to describe how social engineering attacks work and take effect.

and creates more attack opportunities. The easy availability of open source intelligence simplifies the information gathering. Specific targets can be carefully selected to craft more creditable and targeted social engineering attacks. A large group of victims can be reached at the same time and some open source tools can be used to launch semi-automated attacks. Technologies such as machine learning and artificial intelligence is likely to make social engineering attacks more efficient and aggressive. Targeted, large-scale, robotic, automated and advanced social engineering attack is becoming possible [1]. Social engineering is evolving to be a serious, universal and persistent security threat.

To protect against social engineering attack, an important work is to understand how it works and takes effect. This paper makes the following contributions.

- An integrative and structural model to describe how social engineering attacks work and take effect.
- Three core entities to get an insight into social engineering attacks.
- 30+ effect mechanisms involving 6 aspects.
- 40+ human vulnerabilities involving 6 aspects.
- Case study of 16 social engineering attack scenarios (including 13 type attack methods).

II. A CONCEPTUAL MODEL OF HOW SOCIAL ENGINEERING ATTACKS WORK AND TAKE EFFECT

In a cyber-attack, attacker and victim (target) are entities at the two ends. For social engineering, the attacker (a.k.a. social engineer) is the party conducting a social engineering attack; the victim is the party suffering a social engineering attack and bring about an attack consequence. In general, the social engineering attack process can be described as follows: 1) the attacker crafts certain attack methods to exploit the target's human vulnerabilities and further to achieve certain attack goals; 2) once the human vulnerabilities exploited, the target turns into a victim and brings about certain attack consequences; 3) the attacker feeds the consequences back to the attack goal, to decide the next actions.

There are three basic perspectives to understand how social engineering attacks take effect, as Figure 1 shows.

- From the attacker perspective, the attack method is the way, manner or means of carrying an attack out; it is also the driving force that directly causes a social engineering attack and significantly affects whether the attack can succeed. After all, the advanced and ingenious attack methods usually possess a greater success rate to obtain the attack goals.
- From the victim perspective, the exploited human vulnerabilities are the root reason why the victim brings about the attack consequences. As one of the confrontational focuses between social engineering attack and defense, human vulnerability is what attackers want to exploit and what victims want to eliminate or mitigate. Other types of vulnerability (e.g. software vulnerabilities) can be exploited together with human vulnerability, yet they are not necessary in social engineering attack [1].
- From the perspective of principle and explaining, effect mechanisms explain how attack methods make the human vulnerabilities take effect. Effect mechanisms describe [R1] how attack methods exploit human vulnerabilities, and explain [R2] why the human vulnerabilities leads to the attack consequences as well as (corresponding to) [R3] how the attack methods achieve the attack goals. In other words, effect mechanisms can be defined as the structural relation that what, why or how specific attack effects consequences) correspond to specific human vulnerabilities, in specific attack scenarios.

Thus, effect mechanism, human vulnerability and attack method can serve as three core entities to get an insight into how social engineering attacks work and take effect.

We will analyze and discuss the effect mechanisms and human vulnerabilities in the Section III and Section IV respectively. Section V will study a set of social engineering attack scenarios where many attack methods are included, to illustrate how these mechanisms, vulnerabilities and attack methods explain the success of social engineering attacks. Section VI shows the discussion. Section VII concludes the paper.

III. EFFECT MECHANISMS IN SOCIAL ENGINEERING

This section analyzes and discusses social engineering effect mechanisms in 6 aspects: 1) persuasion, 2) social influence, 3) cognition, attitude and behavior, 4) trust and deception, 5) language, thought and decision, 6) emotion and decision-making.

A. EFFECT MECHANISMS IN ASPECT OF PERSUASION

1) SIMILARITY, LIKING AND HELPING IN PERSUASION

Similarity invites liking, dissimilarity leads to dislike. The more someone's attitudes are similar to our own, the more we will like the person [2]. On the contrary, we tend to decrease liking when getting to know someone and discovering the person is actually dissimilar [3]. Furthermore, similarity is conducive to liking and liking elicits helping. We are more empathic and helpful toward those similar to us [4]. ''We most prefer to say yes to the request from people we know and like'' [5]. Besides, physical attractiveness also influences our willingness to help. Attractive people receive more help than those unattractive [6], [7].

Thus, it may be less effective that a social engineer (attacker) attempts to over-persuade the targets in a manner obviously against their inclination or thought. The conflicts of opinion and attitude not only lead to targets' dislike, but also may imply that you are more intelligent than them, which cause the feeling of discontent. In this situation, the persuasion is difficult. The less conflict with the target, the better. Cooperation will be more readily gained when a soft approach is used [8]. It is a good choice facilitating a successful persuasion in social engineering that pretending to be a person who shares the targets' ideas, who likes the same things, or who has a spatio-temporal proximity.

2) DISTRACTION IN PERSUASION AND MANIPULATION

People typically have a limited range of attention in sight, hearing and thought. Distraction facilitates persuasion mainly by disrupting the counter-argue process and increasing the effort to communication. It is effective both online and on the scene. Distraction may force the target to exert high effort so that to hear and understand the persuasive message. Experiments show that moderate distraction does facilitate persuasion, and moderate distraction produces more persuasion than strong distraction because targets are less inclined to suspect the persuasion is intended [9]. The present distraction increases participants' yielding to propaganda by inhibiting counter arguing [10]. Online advertises that frequently disrupt people's web surfing actually do have a persuasive effect even when people do not actively attend to them [11]. ''Although consumers maintain illusory beliefs that they can tune out such ads, the ads have substantial persuasive and subtle distracting effects'' [12]. Distracted persons who have a low propensity to counter argue will be the least resistant to persuasion [13]. Distraction is often used in malicious manipulation attacks. The thought process regarding security will be inhibited and disrupted if the target's focus is transferred to elsewhere.

3) SOURCE CREDIBILITY AND OBEY TO AUTHORITY IN PERSUASION

People have a tendency to comply with authoritative figures automatically. In most cultures, especially the collectivist culture, people are told that to believe who are authoritative, expert and familiar, since these characteristics signify the credibility, trustworthiness and low-risk. For individuals low in need for cognition, when the message source was assumed to be relatively honest, persuasion are less dependent on message scrutiny [14]. Experiments on obedience to authority show that, authority is so powerful that our independent thinking and rational behavior are often suppressed [15], [16]. Even the symbols of authority can trigger the individual's compliance. For instance, in the experiment conducted by study [15], the hospital nurses were ordered by an unknown physician (who stands for expert and authority) to administer patients an obvious overdose drug. Although almost all the nurses and nursing students in the control group claimed they would not to obey, in the experimental group, all 22 nurses but one obeyed without a delay despite the order was given by the phone, until they were intercepted on their way to the patients.

This explains why the symbols that reflect the authority, expert and credibility, such as uniform, badge, lingo and insider terminology, are frequently used in social engineering attacks [17]. Study [18] also shows that authority is effective to convince targets that the phishing URLs in emails are secure.

4) COGNITIVE RESPONSE MODEL, TWO ROUTES TO PERSUASION AND ELABORATION LIKELIHOOD MODEL

Petty [19] conducted a cognitive response analysis of the persistence of attitude changes induced by persuasive communications, in which a cognitive response model was proposed to show that enduring attitudes changes are the result of cognitively responding to the message content, while temporary attitudes shifts are the result of persuasion cues. Cognitive response occurred to thoughtfully process the communication when recipients have both the motivation and the ability. If a recipient is motivated (issue involvement, relevance, commitment, source credibility, etc.) and has the ability (e.g. message is not extremely complex, unfamiliar, fear appeals) to process the content, a change in cognitive structure will lead to an enduring attitude change; otherwise, perceived persuasion cues will be processed. This was later developed into the central and peripheral routes to persuasion [20]. Based on the study of two routes to persuasion, Petty and Cacioppo [21], [22] developed the elaboration likelihood model to discuss a wide variety of variables that proved instrumental in affecting the elaboration likelihood, and thus the routes to persuasion.

The central route occurs when targets motivated by some factors and have the ability to think about the issue, in which the targets to be persuaded are in a high involvement state and the arguments are examined and processed elaborately. The motivation may be things interesting, important or personal related. The target is very likely to be persuaded if the arguments are strong and compelling; while if only weak arguments are perceived, thoughtful people will counter argue. Usually, explicit and stable attitude changes can be obtained when people persuaded in the central route. On the other hand, when the targets are not able to think carefully or motivated (e.g. they are busy or distracted, the message is boring), they might follow the peripheral route of persuasion. In this situation, the targets to be persuaded are in a low involvement state; there may be no ability or motivation to analyze argument quality and reflect on the substance of messages, and the arguments will not be concerned to process elaborately. The cues that trigger automatic acceptance without much thinking will be perceived; the statements which are familiar or easy to understand will be more persuasive than the novel statements with the same meaning; messages with greater number of arguments will be more persuasive than those with fewer arguments.

Individual difference also occurs in persuasion, e.g. recipient's cognitive propensity affects the persuasion. Need for cognition refers to an individual's tendency to engage in effortful cognitive activity and enjoy cognitive endeavors [23]. Individuals low in need for cognition often act as cognitive misers [24] and peripheral cues that rarely required cognitive endeavors are perceived swiftly. In contrast, individuals high in need for cognition are more likely to think about issue-relevant information and prefer the central route where message are processed elaborately and cognitively. Similarly, for individuals high in need for cognition, message scrutiny did not differ depending on the source [14].

For social engineering attacks, people who take computers as essential work tools are highly involved, such as system administrators, computer security officers and technicians [8]. These targets are more likely persuaded by strong arguments with central route, and weak arguments tend to generate challenge. People such as security guards, cleaners and receptionists are considered as lowly involved [8]. They are generally not able to understand the technical context or have little interest in the request contents. They tend to avoid bother analyzing the request deeply and make decision based on the peripheral cues. This paper considers that persuasion is not an "either / or" choice, but a "more and less" way. A recipient can process arguments thoughtfully, meanwhile influenced by the peripheral cues. Peripheral processing also requires some cognition, and a target can be persuaded by two routes simultaneously. Therefore, although the peripheral route is frequently used in social engineering [25], [26], as long as the pretext is credible and compelling, the central route will also take effect.

B. EFFECT MECHANISMS IN ASPECT OF SOCIAL INFLUENCE

1) GROUP INFLUENCE AND CONFORMITY

People live in and influenced by groups almost all the time. Conformity is a change in behavior or belief to accord with others as the result of real or imagined group influence [27]. There are many factors affect the conformity, such as group size, group unanimity, group cohesion and individual's public response. A small group can lead to a big conformity effect. People conform distinctly when the group increased to a certain size. As group size increased from 3 to 5, the percentage of passersby who imitated a group looking upward increased from 60% to 80% [28]. If the group unanimity of behavior or belief decreases, the conformity will reduce. Group cohesion enhances the conformity, and people also conform most when their responses are public [27].

2) NORMATIVE AND INFORMATIONAL INFLUENCE

Usually, an individual may bend to the group in order to be accepted or to obtain important information. The former is called normative influence and the latter is called informational influence [29]. Conformity caused by normative influence is motivated by the desire to be accepted or liked, or to avoid group pressure. When deviating from social group norms, people often bear social pressure and pay an emotional price. After all, for most people, social rejection is painful. Thus, individuals conform with groups intentionally or unintentionally to seek groups' acceptance and appreciation, in which smoking, drinking, alcohol and drug abuse, steal and other dangerous actions occurred [27]. Here, the correctness is not that important. This is also known as social validation. In the conformity due to informational influence, people attempt to find the correct decision or avoid unknown risks. People typically assume the group actions are more right and less risk, which influence them to adopt the same behavior, belief and decision. This is also known as social proof. "We determine what is correct by finding out what other people think is correct [5]". Conforming with groups may be beneficial in some situations, however, accepting informational influence without thinking will lead to a blind follow.

In social engineering attacks, the attacker often craft specific information and scenario, in which normative influence and informational influence are used to influence and manipulate the targets to do certain actions that benefit the attacker.

3) SOCIAL EXCHANGE THEORY AND RECIPROCITY NORM

Social exchange theory shows that people exchange not only material goods and money but also social goods such as love, services, information and status [30]. The consideration or subtle calculation about cost and reward predict people's decision and behavior (e.g. help). Reciprocity norm refers that we should return help but harm to those who help us [31]. We shall try to repay similar with what another person has provided us. If others do us a favor, we shall do them a favor in return [5]. As a universal social norm, reciprocity always influence people along with the process of socialization. People universally internalize the idea that reciprocating others for their kindness and help. Besides, for all social interactions, the exchange ought to be balance in the long term. Receive without giving in return violates the reciprocity norm.

In a reverse social engineering attack, the attacker impersonates a person who belongs to system administrator, IT department, help desk or technical support, and then wait (e.g. a new employee) to ask a help to solve a computer or network fault. Once this occurred, the attacker attempts to exploit the new employee by requesting a favor or eliciting a password. It succeeds because the new employee is expected to reciprocate the attacker in the security context and fulfill a social exchange.

4) SOCIAL RESPONSIBILITY NORM AND MORAL DUTY

Different from the reciprocity norm where the balance of giving and receiving are considered, social responsibility norm advocates that people should help those who need help, without concerning the future reciprocate and exchanges [32]. It is a kind of expectation towards moral duty for helping. In collectivist culture countries, people support the social responsibility norm more strongly than individualist culture countries [33]. They advocate an obligation to help others even they are not facing a life-threatening trouble.

Social responsibility norm and moral duty take social engineering attack effect by at least two ways. One way is that the attacker exploits the targets' tendency to be helpful (which internalized in the forming of social norm) to elicit information or obtain a favor facilitating the attack. Another way is that during the social engineering attack, the group pressure caused by social responsibility norm and moral duty is used to influence targets' behavior, especially for the targets who are not willing to provide a help.

5) SELF-DISCLOSURE AND RAPPORT RELATION BUILDING

Derlaga and Berg [34] researched on the self-disclosure and described the disclosure reciprocity effect. It shows that during the building of social relation, self-disclosure begets self-disclosure, and we have a willing to reveal more to those who open their hearts to us. It is gratifying to be selected as the person for another's self-disclosure. Not only do we like those who open their mind to us, we also disclose to those whom we like, and we like them more after disclosing to them [35]. Some people (most of them are women) are particularly skillful at getting people to open up; they can easily elicit intimate disclosures from others, even from those who normally don't reveal very much of themselves [36]. Disclosure that open up to another person, just like taking off our masks and letting ourselves be known as we are, nurtures the rapport relation [37], implies trust and facilitates the social interaction.

C. EFFECT MECHANISMS IN ASPECTS OF COGNITION, ATTITUDE AND BEHAVIOR

1) IMPRESSION MANAGEMENT, COGNITIVE DISSONANCE AND COMMITMENT AND CONSISTENCY

It is a human nature to care about what others think of us. Self-presentation theory shows that we want to present a favorable impression both internal to ourselves and external to other people, so that to feel better about ourselves, to gain social and material rewards, and even to become more secure in our social identities [27], [38]. We make great effort to manage our behaviors to create a desired image. In order to keep our creditability and protect our self-esteem, no one wants to look inconsistent. Thus, we manage our image by behaving in line with attitudes or commitments we have presented, or by expressing attitudes that match our actions.

Cognitive dissonance theory [39] shows that we feel tension or a lack of harmony (dissonance) when two psychologically inconsistent cognition (thoughts, beliefs, etc.) are simultaneously perceived. And to reduce this discomfort, we often adjust our thinking, especially when external inducements are insufficient to justify our behavior. Cognitive dissonance occurs, e.g. when faced with an important decision between two equally attractive alternatives, 1) we subjectively make one selection although reasons support another, or 2) we recall the advantages of what has rejected and the disadvantages what has chosen. To reduce the cognitive dissonance, we may justify our selection by adjusting our ideas and even revising our memories, or we match the results with our cognition by behaving to change the selection.

These two theories explain the unity of opposite between behavior and attitudes from different perspective, and also explain the underlying mechanism of commitment and consistency [5], i.e. once we made a choice or took a stand, we experience supervision and pressures from both inside and outside, forcing our words and behaviors consistent with what we have commitment.

2) FOOT-IN-THE-DOOR: BEHAVIOR AFFECTS ATTITUDE

If you want people to do you a big favor, an effective strategy is to get them to do a small favor first. In an experiment, experimenters who claim they are from the Community Committee for Traffic Safety asked some Californians (control group) to install a very large sign that said "Drive Carefully" in their front lawn; only 17% people consented. Another group of people (experimental group) were first approached with a small request that whether they agreed to take a small (three-inch square) sign of "Be a safe driver" and put it in a window or in the car so that it would serve as a reminder of the need to drive carefully; nearly all readily agreed; and when contacted two weeks later to permit the large and poorly lettered signs in their front lawn, 76% people consented [40]. This phenomenon that once someone has agreed to a small request, he or she is more likely to comply with a larger request is known as the foot-in-the-door effect. This effect is still effective in online context as well as in other communication modalities, since it functions through an individual's internal consistency motives [11]. Compared to attitude predicts behavior, the foot-in-the-door effect shows that behavior affects attitude. It seems that people build a image after they did the small favor (behavior). In order to maintain the consistency of this image or relieve the pressure or cognitive dissonance caused by differences between internal attitude and external behavior, people attempt to manage their later attitudes and behaviors to be consistent with their previous behaviors. This also explains why helpfulness is frequently exploited in social engineering attacks.

3) BYSTANDER EFFECT, DIFFUSION OF RESPONSIBILITY AND DEINDIVIDUATION

Bystander effect describes the phenomenon that a person is less likely to provide help when there are bystanders' presence. In other words, the person who needs help is actually less likely to get help when many people are around [41]. The person in need is more likely to get help when bystanders present alone, and the more bystanders to an emergency, the less likely or the more slowly a bystander will intervene to provide aid [42], [43]. In large cities, the increasing numbers of bystanders who are strangers often depress helping [44]. Bystander effect reflects the phenomenon of diffusion of responsibility in the context of helping. The presence of other bystanders reduces the individual's feeling of personal responsibility for helping [43]. The same is true in online environment. Those who received the email (requesting help) along with an indication that no or few other people were also contacted, provide significantly more assistance than those who received the request along with an indication that many others were also contacted [45].

When individuals in groups abandon normal restraints, forget their individual identity and follow the group or crowd norms, deindividuation [46] occurred. It's in group situations that people are facilitated to lose their self-awareness and evaluation apprehension, to diffuse responsibility across all group members, and to respond to group norms regardless of its good or bad. Deindividuated people perceive their actions as the group's. Anonymity, large group size, arousing and distracting activities are factors promote the deindividuation. The bigger the mob, or the more anonymity (e.g. night, face mask), the more its members lose self-awareness and become willing to commit atrocities. When people are not accountable and their own behavior can not be evaluated, bystander effect, diffusion of responsibility, deindividuation and social loafing [47] occurred. In social engineering attacks, victims may be induced into certain group situations and exploited by these effect mechanisms to perform actions that endanger cyberspace security (e.g. case 16 in Table 1).

4) SCARCITY: PERCEIVED VALUE AND EMOTION-AROUSING

Scarcity manipulates people mainly by affecting value cognition, arousing emotion and enhancing motivation. "Opportunities seem more valuable to us when they are less available" [5]. Economics and social experience told people that the scarce resource implies less accessible, more competing risk and less freedom. Hence, people assign more value to the scarce things, although usually this subjective value are overestimated. To avoid this potential risk and respond to the perceived value, people have a stronger desire than before to gain this resource, even though the scarcity may be temporary, unnecessary or even faked. Moreover, scarcity enhance the motivation and drive the behavior by arousing emotions such as fear, anxiety, desire and greed. A fear-arousing message is potent when persuading people to cut down on smoking, drive carefully or get a tetanus shot [27]. In many attack scenarios, attackers exploit victims to disclosure information or trigger a malicious URL by convincing them what offered is scarce (e.g. phishing in Table 1).

5) TIME PRESSURE AND THOUGHT OVERLOADING

Time pressure affects people's logical thinking. When people have to deal with a large amount of information in a limited time, request messages that shall be examined are often responded rashly and superficially. Besides, time pressure might lead to emotion-arousing, such as anger, tension and anxiety, which inhibits cognition by making thinking difficult [5]. Complex messages or non-grammatical sentences may lead to a thought overloading. An example can be that "Do you realize that you're not thinking right now of what I am not saying? And can you realize that it's not that easy to not know what I am going to say next, but even when you're not knowing it I am knowing it and you're not?" [48] Our brain shuts down when faced with similar messages. This overloading of thought is similar to the overloading of computer, such as denial of service caused by running out of memory and buffer overflow attack. A thought overflow attack seems feasible given that the target's thought can be exploited like the computer buffer: specific return / jump address (thought anchor) are overwritten / redirected to the malicious payload by the overflowed buffer (overloaded thought). So, attackers might dedicate to submerge the targets with lots of misleading information and create them a sense of urgency, in order to trigger their thought overloading, disrupt their normal cognizing or thinking, and further elicit a vulnerable behavior.

D. EFFECT MECHANISMS IN ASPECTS OF TRUST AND DECEPTION

1) RELATION BETWEEN TRUST AND SOCIAL ENGINEERING

Trust is an important variable that predicts the user's susceptibility to social engineering attacks [49]. Chitrey *et al.* [50] conducted a survey showing that "90% of the participants think that people in India generally have a higher level of social trust, which implies that they are more vulnerable to social-engineering based attacks". In many social engineering attack scenarios, it requires to convince the targets that the attacker is a trustworthy person [51]. According to [52], trust is the willingness of a party (trustor) to be vulnerable to the actions of another party (trustee) based on the expectation that the other (trustee) will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party (trustee). That is, trusting behavior is essentially a risk taking. Thus, the positive correlation between trust and the low risk of willingness *to help, build rapport relation and share information*, may be the significant reason why attackers pay close attention to trust. In some attack scenarios, even though certain risk have been perceived, social engineering attacks still occurred because of a stronger trust has been built. Therefore, factors affecting trust building and factors affecting deception become the parameters that attackers want to control, to perform a social engineering attack.

2) FACTORS AFFECTING TRUST

There are three basic objects involved in analyzing factors affecting trust building: the trustee (attacker), the trustor (target, victim) and situation. Mayer *et al.* [52] presented an integrative model of organizational trust, in which trust propensity, perceived trustworthiness of trustee and perceived risk are considered as factors affecting the trust behavior (a risk taking) of a trustor. Trust propensity describes the general willingness of a trustor to trust others previous the knowledge about them. It is a stable factor that affect the likelihood the trustor will trust. People differ in their inherent propensity to trust for their different developmental experiences, personality types, and cultural backgrounds. Three characteristics of the trustee (ability, benevolence and integrity) affect the perceived trustworthiness. Trust is developed based on both trustor's propensity and perceived trustee's trustworthiness. Trustor takes trusting behavior in a relationship under the consideration of perceived risk; if the outcomes of trusting behavior are positive, the trustor's perceptions toward trustee will be enhanced. Sztompka [53] discussed other factors affecting trust building, such as reputation, performance, appearance, accountability and trust-inducing situation. Besides the above factors, other situational factors might also affect the trust building and the success of social engineering attacks, such as cyber-environment (instant communication, social network, websites, etc.), social culture, security strategy and natural environment. For social engineering attacks, the target's characteristics including the trust propensity are factors the attacker can identify but can little control. Therefore, attackers usually try their best to exhibit their trustworthy factors and manipulate the situational factors to exploit the target's human factors about trust.

3) FACTORS AFFECTING DECEPTION

Usually, deception is intentional, strategic interaction behaviors launched by the deceiver. Although most people are confident that they can detect social deception, interpersonal deception theory (IDT) [54] suggests that they cannot. IDT attempts to explain the process and outcomes of deception in interpersonal conversations based on the deception analysis, propositions and evaluation. According to IDT, in deception communication, senders take more strategic activities to manage / manipulate information, behavior and image to deceive receivers; reciprocity is the predominant interaction adaptation pattern between senders and receivers [54]; during interpersonal deception, three types of deception that falsification (e.g. lie, create a fiction), concealment (e.g. part truth, hides a secret) and equivocation (e.g. dodges the issue) are frequently used. IDT discussed many factors that affect the deception and its detection: 1) increased interaction produces greater strategic activities; 2) expectations for honesty inversely related to deceivers' fear of detection and associated strategic activity; 3) the more skilled senders are, the more truthful demeanor are conveyed and the less nonstrategic leakage; 4) receivers' suspicion and detection is positively related to receivers' truth biases, context interactivity, senders' encoding skills, informational and behavioral familiarity, receivers' decoding skills, and senders' deviation from expected patterns [54].

E. EFFECT MECHANISMS IN ASPECTS OF LANGUAGE, THOUGHT AND DECISION

1) RELATION BETWEEN LANGUAGE AND THINKING

Language is the most common tool for social interaction meanwhile it is closely related to the processing, generating and expressing of thought. Language can be compared to the computer program used for communication. The words we hear are the inputs and the streams of thought are outputs, vice versa. We either listen to what others are saying and attempt to visualize, or we try to put what we are thinking or visualizing into words [48]. It is language that enables people to encode everything perceived into operable symbols in the brain and further to think about them. Almost everyone attempts to seek appropriate words to express ideas and feelings, especially when these ideas and feelings are vague. In this sense, thinking and social interaction depend heavily on language. For social engineering attacks, elaborately crafting the information conveyed to the receiver (inputs of language) implies manipulating the receiver's thinking and the corresponding behaviors, i.e. the language cognition is exploited.

2) FRAMING EFFECT AND COGNITIVE BIAS

Framing effect is an interesting phenomenon reflecting cognitive bias, in which people make decisions and express opinions influenced by the way a question or an issue is described. In other words, for the same problem with different expression, different choices are made. For instance, beef labeled as "25% fat" versus beef labeled as "75% lean", the latter is preferred usually. Another case in point is that, when asked that "should the government allow this type of condom to be advertised an effective method for reducing the risk of AIDS?"; if the condom described as "95% success rate", 88% participant students selected "Yes"; however, if told its "5% failure rate", the selection of "Yes" down to 42%, i.e. 58% participants selected "No" [55]. This cognitive bias for language framing leads to a decision manipulating. Thus, when the targets are induced into a predesigned language framing through which a cognitive bias is exploited, a successful persuasion or decision manipulating for social engineering attack is likely to be achieved (e.g. case 16 in Table 1).

3) INDIRECTNESS OF THINKING AND NEGATIVE EXPRESSION IN LANGUAGE

The dependence of thinking on language (Section III-E1) leads to the indirectness of semantics transmitting, which creates opportunities for language hinting and inducing. Furthermore, the cognitive indirectness for negative language expressions can also result in influence and manipulation. For example, "If I say the car is not driving to the store. How do you visualize that? You can only visualize the action then wipe it away" [48]. Another case in point can be that "there is not a book on the desktop", for which, we create in mind firstly a book and then a desktop, finally we take the book away from the desktop. When received negative language expressions, our mind firstly constructs the positive concepts, imagery or actions for the semantic units and then negate them. Although some of these concepts, imagery and actions may be insignificant, what's significant for social engineering is that they do occur in the mind and influence the decision-making, no matter what the semantic meaning of the sentence itself is.

4) LANGUAGE EVOKES THINKING CONFUSION

Language can be used to evoke a thinking confusion state, in which behaviors are suggested and commands are embedded; this provides the attacker an opportunity to induce and manipulate the targets to take actions that may breach security policy. A simple method to evoke thinking confusion is expressing with non-grammatical or ambiguous statements. This force the targets to search inwards about what the attacker just said meanwhile trying to concentrate on what the attacker is currently saying, as a consequence the critical faculty of normal thinking is bypassed. The expression like "that's touching to hear" may result in the targets touching their ear [56].

F. EFFECT MECHANISMS IN ASPECTS OF EMOTION AND DECISION-MAKING

1) EMOTION AND FEELING AFFECT DECISION-MAKING

A familiar view regarding human decision-making is that people make decisions through the dual systems of emotion and reason: one is generally emotional, fast, automatic, and the other is cognitive, slow, deliberative. In fact, the mechanisms of emotion and decision is very complex. Research related to emotion and decision-making has achieved many results, such as emotion's limbic system theory, emotional brain (amygdala) and two circuits (subcortical circuit and cortical circuit) theory [57]. Phelps *et al.* [58] shows that there are multiple neural circuits underlying the modulation of decision-making by emotion or affect. There are many challenges in emotion and feeling research, e.g. it is difficult to accurately manipulate and measure emotion and affect. However, this field evolves along with the development of cognitive science, neuroscience, brain science, anatomy and instrument techniques [58]. Although certain incompatibility occurred, one thing agreed by numerous studies in different areas(e.g. social psychology, neuroscience, brain science and anatomy) is that emotion and feeling do affect the decision-making. Emotion and mood elicit the action tendencies and carryover it onto the decision process [58]. Individuals in sad emotion tend to seek high-risk / high-reward options, but individuals in anxious emotion prefer the low-risk / low-reward options [59]. People in fear emotion express pessimistic risk estimates and risk-averse choices, yet people in angry emotion express optimistic risk estimates and risk-seeking choices [60]. In a happy state, people are more helpful, and decisions are more quickly [61]. Positive mood can lead to more persuasion [62].

For social engineering attacks, it implies that manipulating emotions and feelings will influence and even alter the targets' decision. Changing emotion can change choices; methods such as cognitive emotion regulation and targeting re-consolidation can be used to modify emotions in different decision contexts [58]. The unconsciousness conditioned emotional response [63] also reminds us to protect our cognition, neuro and emotion from being exploited.

2) EMOTION, FACIAL EXPRESSION, DECEPTION AND DECEPTION DETECTION

For social engineering attacks where deception is used, the attacker as the deceiver will pay greater cognitive exertion to exhibit strategic information, behavior and image management meanwhile strive to avoid nonstrategic deception leakage. However, with the increasing of receivers' familiarity towards information, behavior and relation, the attacker not only experience more detection apprehension but also exhibit more nonstrategic leakage behavior [64]. The leakage of deception is usually reflected on non-verbal signals, especially facial expressions. Non-verbal signals permeate in the vast majority of social interactions and people perceive and comprehend them consciously or unconsciously. The non-strategic leakage of verbal deception occurs when non-verbal signals that reflect the true intention are not inconsistent with the deceptive verbal signals. Facial expressions may be the most directly external non-verbal signals that reflect the internal true intentions and emotions. Thus, it is also an arena for social engineering, in which the attacker attempts to deceive the targets by inhibiting deception leakage whereas the victim or defender strives to detect the deceptive attack by identifying the nonstrategic leakage of facial expressions.

A micro expression is a momentary involuntary facial expression as a result of a voluntary and an involuntary emotional response occurring simultaneously and inconsistently. Although the micro expression disappears quickly, it reflects a person's true emotions. Thus, micro expressions serve as a set of nonstrategic leakage offers a way for emotion recognition and deception detection. Ekman *et al.* [65] updated the Facial Action Coding System (FACS), in which various kinds of facial muscle movements are encoded as a series of action units which further are combined to describe different emotions. Micro Expression Training Tool (METT) and Subtle Expression Training Tool (SETT) [66] have been also developed for facial expression recognition analysis and training. These tools related facial expressions and micro-expression are helpful in social engineering defense.

IV. HUMAN VULNERABILITIES IN SOCIAL ENGINEERING

Social engineering attacks exploit a wide range of human vulnerabilities. This section discusses these vulnerabilities in the following aspects: 1) cognition and knowledge, 2) behavior and habit, 3) emotion and feeling and 4) psychological factor. And the psychological vulnerabilities are further divided into three levels, i.e. 1) human nature, 2) personality traits and 3) individual characteristics, from the evolution perspective of human wholeness to individuation.

A. HUMAN VULNERABILITIES IN COGNITION AND KNOWLEDGE

Thinking set (inertial thinking) is a relatively rigid way, process or mode to think about something. It can be also described as a relatively stable behavioral tendency or psychological readiness state that derived from / built on the previous experience and cognition. Thinking set helps people quickly address problems in the familiar environments, yet it will hamper the right treatment to new matters when situation changed. Stereotype and prejudice are similar vulnerabilities.

Cognition based on heuristics or mental shortcuts, e.g. intuitive and impulsive judgements, are more likely to be exploited by persuasive social engineering attempts [67]. For employees who are indifferent to their work, it will be very difficult to ensure information security, especially when shortcuts are taken and security rules are not followed [68].

Conformity (Section III-B1) and low level of need for cognition (Section III-A4) are vulnerable to persuading, influencing and manipulating in social engineering attacks. Ignorance and conformity are susceptible to a social engineering attack [51]. Ignorance (e.g. low awareness of information value and security) is easily exploited by attackers, e.g. through "direct approach [69]". Ignorance and inexperience bring security risk [70]. Most of the mobile users in Kenyans fall prey to vishing attack due to the lack of knowledge on social engineering [71]. Users should be educated on what constitutes sensitive information and how it can be abused in online attacks [72].

B. HUMAN VULNERABILITIES IN BEHAVIOR AND HABIT

When a person does not pay enough attention to the security context (carelessness [69]), does not think about the potential security risk (thoughtlessness [73]), or is unwilling to make necessary work or effort to prevent a security threat (laziness [74]), the person will be a target through whom a social engineering attack occurs easily.

Fixed action pattern exists in behaviors of both animals and humans, which consists of a series of relatively invariant instinctive behaviors triggered by a key stimulus. This set of actions are automatic and involuntary, and will go to completion even if the stimulus is removed. The behavioral habits are similar to the fixed action pattern, yet they are largely voluntary and automatic. Owing to these behaviors are voluntary, automatic and subconscious, they are vulnerable to social engineering attacks, and the targets are hard to aware that they have been exploited. For instance, in the water-holing attack, if the targets have a habit to periodically visit certain websites, the attacker may infect these websites with malicious code in advance and wait the targets to visit and trigger.

C. HUMAN VULNERABILITIES IN EMOTION AND FEELING

Emotions and feelings influence cognition, attitude and decision-making (Section III-F1, III-C). Emotions (fear, tension, curiosity, excitement, surprise, anger, impulsion, etc.) and feelings (happiness, sadness, disgust, guilt, etc.) are all human factors can be exploited as security vulnerabilities in social engineering attacks. Fear of getting into trouble with the superiors is often used in name-dropping approach to elicit sensitive information, and fear-arousing presented in Section III-C4 is also a case in point. When strong emotions such as anger, excitement, fear or anxiety are triggered, an individual's cognitive ability may be seriously hampered [75]. Curiosity is vulnerable to baiting attacks. A USB baiting experiment [76] showed that 15/20 USB drives with Trojan functions were found by employees and all had been plugged into company computers. Angry (Section III-F1) and impulsion lead to more risk taking. Users who are less impulsive in making decisions generally are more likely to judge a link in a fraudulent email as unsafe [18]. Guilt is often accompanied by shame and regret, and all these are feelings people try to evade. These feelings turn into vulnerabilities in social engineering, e.g. when attackers managed to convince the target that they will suffer greatly (e.g. scolded by the boss and get fired) if the request is not granted. Here, a foreseeable feeling of guilt stimulates the target to make a softhearted but security-breached decision.

D. HUMAN VULNERABILITIES IN HUMAN NATURE

Human nature is a collection of psychological characteristics at the macro level, which describes the fundamental psychological characteristics shared naturally by the whole human being. Some human natures are security vulnerabilities exploitable in social engineering attacks. People who pay close attention to themselves and their desires will magnify the ambient influence and increase the susceptibility to induce, persuade and manipulate in social engineering. Unbridled demand for lures and the fear of being excluded from potential gains will drive weak-willed people into vulnerable decisions or risky behaviors. Thus, human natures such as self-love (narcissism), greed, lust and gluttony become vulnerabilities that can be exploited in certain social engineering attack scenarios [70], [73], [77]. People naturally sympathize with individuals who are in trouble.

Sympathy and helpfulness are universally recognized virtues in different social cultures, yet bring potential risks in the context of cybersecurity. Pretending to be a person who needs help proves effective over and over again in allowing social engineers to reach their goals [17].

E. HUMAN VULNERABILITIES IN PERSONALITY TRAIT

Individuals' personality traits significantly contribute to their susceptibility to social engineering exploits such as influence, manipulation and deception [78], [79]. Social engineers treat human personality traits as vulnerabilities and use the language as their weapon to deceive, persuade and finally manipulate the victims [80]. Personality traits are the psychological structure or characteristic set of habitual patterns of behavior, thought, and emotion, which evolve from the biological inheritance predominantly with the influence of environmental factors. Personality traits differ across individuals and influence their behaviors. Besides, personality traits are relatively stable over time and consistent over situations. In many theories and systems, personality traits are classified to different dimensions in which an individual's traits can be rated along the spectrum. The five-factor model of personality is a hierarchical organization of personality traits in terms of five basic dimensions, i.e. extraversion, conscientiousness, agreeableness, openness to experience, neuroticism or emotional stability [81].

Personality traits in the dimension of extraversion manifest mainly as activity, warmth, positive emotions, assertiveness, excitement seeking and gregariousness. Individuals with high extraversion are more active, enthusiastic, assertive, energetic, outgoing and talkative. Thus, they are vulnerable to social engineering by effect mechanisms such as similarly & liking & helping, self-disclosure and rapport relation building, impression management, commitment and consistency, risk taking for trust, and conformity.

Conscientiousness dimension concentrates on competence, order, dutifulness, self-discipline, achievement striving and deliberation. People in this kind are more efficient, organized, responsible, planful, reliable and thorough. Thus, they are vulnerable to social engineering by effect mechanisms such as central route of persuasion, obeying to authority, informational influence, social responsibility norm, moral duty, and commitment and consistency.

Personality traits in dimension of agreeableness incorporates trust, straightforwardness, altruism, compliance, modesty and tender mindedness. People with more traits of agreeableness are trusting, appreciative, generous, sympathetic, forgiving and kind. Thus, they do more credulous actions and are vulnerable to social engineering by effect mechanisms such as group influence and conformity, social validation, reciprocity norm and foot-in-the-door.

Dimension of openness to experience pay attention to fantasy, aesthetics, feelings, ideas, values and actions. People belong to this type is more imaginative, artistic, curious, insightful, original and wide interests. Thus, they are vulnerable to social engineering by effect mechanisms such as peripheral route of persuasion, various kinds of emotion-arousing and cognitive dissonance.

Personality traits in neuroticism dimension include anxiety, hostility, self-consciousness, depression, impulsiveness and vulnerability. People with high neuroticism is more anxious, tense, worrying, self-pitying, unstable and touchy. Thus, they are vulnerable to social engineering by effect mechanisms such as fear-arousing, cognitive dissonance, evaluation apprehension, diffusion of responsibility and deindividuation.

F. HUMAN VULNERABILITIES IN INDIVIDUAL CHARACTER

Individual characters are psychological characteristics that acquired with the influence of external environment and developed based on human nature and personality traits. In the context of cybersecurity, when some positive individual characteristics are immoderate or in an inappropriate situation, negative results can be generated. If trust is substituted by credulity, deception occurs easily. Friendliness implies a rapport relation and more disclosure. Kindness and charity may lead the victims to offer more help for attackers. When the unauthorized attacker attempts to enter areas that need access card, humility and courtesy are usually exploited, e.g. the victim will hold a door open for others or let others enter first. Similarly, some negative individual characteristics can also be exploited as security vulnerabilities in social engineering. Diffident people are more likely to obey to the authority and less likely to challenge the attacker's request. People in hubris may disdain to comply security policies and indifferent people may have no interest in or enthusiasm about security risk. Envy can lead to phishing attack by a lure. Thus, individual characteristics such as credulity, friendliness, kindness, charity, humility, courtesy, diffidence, apathy, indifferent, hubris and envy become vulnerabilities in social engineering.

V. CASE STUDY: SOCIAL ENGINEERING ATTACK SCENARIOS ANALYSIS

This section presents 16 social engineering attack scenarios (Table 1) to illustrate how to use the three core entities (i.e. effect mechanisms, human vulnerabilities and attack methods) of the conceptual model to get an insight into social engineering attacks. Some of these attack scenarios are based on cases in work [1], and 13 types of social engineering attack methods are included in these 16 scenarios.

In Table 1, the first column describes the attack method and scenario, and the 2nd and 3rd column respectively show the corresponding effect mechanisms and human vulnerabilities. These items in the latter two columns cover almost all the effect mechanisms discussed in Section III and the human vulnerabilities discussed in Section IV.

We intended to detail every attack scenario in Table 1, yet in order to avoid generating a set of dangerous attack guide or script, as well as to avoid the verbose caused by the same description or the well-known explanation, a trade-off

TABLE 1. Cases study of 16 social engineering attack scenarios to illustrate the application of the conceptual model.

No.	Social Engineering Scenarios Description	Effect Mechanisms	Human Vulnerabilities
1	**Pretexting**. The attacker requests classified information by pretending to be a cable splicer and pretexting that he is wiring two hundred pair terminals for police. Who would want to refuse a little help to a company man coping with that heavy-duty assignment? She feels sorry for him, she's had bad days on the job herself, and she'll bend the rules a little to help out a fellow employee with a problem.	Social responsibility norm and moral duty, similarity & liking & helping, emotions and feelings influence decision-making, ELM, IDT, factors affecting trust.	Sadness, sympathy, the desire to be helpful, agreeableness, kindness and charity.
2	**Vishing and Pretexting**. The attacker pretends to be a new employee and convince the targets that he will suffer greatly if the request is not granted. E.g. request the technical support (e.g. Paul) to reset the password of certain account to deal with an urgent task, and further ask a VPN to access from outside.	Foot-in-the-door, impression management theory, ELM, two routes to persuasion, IDT, cognitive dissonance, emotions and feelings influence decision-making.	Guilt, the desire to be helpful, friendliness, credulity.
3	**Vishing and Pretexting**. The attacker calls a staff of the technical support department to say that the CEO authorized his requesting an urgent VPN channel for a project presentation in another city, and further tells he / she that other staffs did this before, such as Paul.	Source credibility and obey to authority, diffusion of responsibility, bystander effect, deindividuation in group.	Fear and dread, neuroticism, the desire to be helpful, friendliness, credulity.
4	**Shoulder surfing**. The attacker pretends to be a delivery man, maintenance worker or consultant to get access to the target workplace and contact with the victims. When the victim is not paying attention, the attacker collects information such as username and password by surfing over the victim's shoulder, snooping prominent places such as sticky notes, papers or computers.	Distraction in persuasion and manipulation, IDT, factor of trust and deception, time pressure and thought overloading.	Carelessness and thoughtlessness, credulity, gullibility, friendliness, ignorance.
5	**Manipulating conversation**. The attackers induce the group conversation to a security topic, one of the attackers discloses his password to discuss whether it is strong enough. If most of the other participants (or attackers) also start disclosing password, the targets are likely to be manipulated to disclose password or other sensitive information.	Group influence and conformity, social validation, reciprocity norm, self-disclosure and rapport relation building, social exchange theory, cognitive dissonance, IDT.	Conformity, agreeableness, extraversion, credulity, courtesy and humility, diffidence.
6	**Piggybacking**. An authorized person provides access to an unauthorized person by keeping the secured door open for providing help or other reasons. Most employees do not know every colleague at a (large) organization and will hold a door open for politeness, let alone the attacker is nicely dressed, shoes shined, hair perfect, with polite manner and a smile; victims will less likely to suspect.	Peripheral route to persuasion, similarity & liking & helping, distraction in persuasion and manipulation, IDT, factors affecting trust, facial expression and deception leakage.	Courtesy, humility, credulity, openness to experience, the desire to be helpful, friendliness, intuitive judgement.
7	**Trailing**. The attacker gaining access to an establishment by following employees who have security card, under the cover of lunch rush at a large corporation. The security guard and employee see in the eye, but he has accustomed to it.	ELM, peripheral route to persuasion, distraction in persuasion and manipulation, level of need for cognition.	Helpfulness, think set and stereotyping, heuristics thinking and mental shortcuts, intuitive judgement, apathy, indifferent.
8	**Trailing**. In some organizations, the lazy security guards put the access card on the desk for those who forget bringing the access card to pick it up for themselves.	ELM, peripheral route to persuasion, level of need for cognition.	Ignorance, lazy and sloth, apathy or indifferent, helpfulness.
9	**Trailing and Impersonating**. The attacker pretends to be an employee of target organization through suitable disguises such as uniform and printed badge, and convinces the janitor of his supposed role, to gain access to a building or a restricted area. When the janitor is the only person present, he is more likely to provide help.	ELM, two routes to persuasion, IDT, factors (appearance) of trust, informational influence, bystander effect.	Think set and stereotyping, intuitive judgement, heuristics thinking and mental shortcuts, credulity and gullibility.
10	**Baiting**. The attacker leaves a USB stick containing malicious codes in a location where it is likely to be found by the victims. The outside of the USB stick is the logo of the target organization or attractive icons to lure the victims to pick up and insert into computer. Once inserted, the malicious code may execute automatically.	Similarity & liking & helping, ELM, two routes to persuasion, IDT, emotions and feelings influence decision-making.	Curiosity, excitement, greed, conscientiousness, sympathy or the desire to be helpful, inexperience.
11	**Phishing**. The attacker sends phishing emails with faked address (or by pup-up windows) to inform targets that there is a very low discount coupons of food (or sport event ticket) in a limited time. The email contains tempting food pictures (or passionate sports posters). This lure the targets to click on malicious links, divulge privacy information, etc.	IDT, peripheral route to persuasion, distraction in persuasion and manipulation, emotions and feelings influence decision-making, scarcity and fear-arousing in persuasion.	Excitement, happiness, greed, gluttony, surprise, extraversion, impulsion, fear, intuitive judgement.
12	**Phishing**. The attacker finds there is some resentment between employees of the target organization through text, images or videos in SNSs, and sends email embedded with malicious code to some of them, claiming it was a hoax virus that could be forwarded anonymously to someone they didn't like. This may compromise a large group of individuals in the organization.	Deindividuation, bystander effect, emotions and feelings influence decision-making, micro expression identifying.	Disgust, prejudice, anger or wrath, hubris, envy.
13	**Smishing**. The attacker blocks the target CEO's cell phone signal and sends SMS message to his secretary by faking the CEO's phone number: "I'm in a meeting at another city and couldn't talk on the phone. Encrypt the organization structure table and a contract file to a zip with key *** and send it to xxx@xxx.xxx immediately! Otherwise, we will lose an important business."	Source credibility and obey to authority, time pressure and thought overloading, fear-arousing in persuasion, emotions and feelings influence decision-making, IDT.	Fear and dread, tension, neuroticism, self-love, credulity.
14	**Trojan attack, honey trap**. The attacker provides a URL and implies it is a porn site, or offering free software (malware) for download to watch porn videos. Text marked that "you won't see the seductive images If you don't act." Once the targets opened the link or installed the software, the attacker's computer or mobile device is compromised.	IDT, emotions and feelings influence decision-making, peripheral route to persuasion, distraction in persuasion and manipulation, indirectness of thinking and negative expression in language.	Lust, greed, excitement, curiosity, impulsion.
15	**Water-holing**. The attacker finds that the targets usually, regularly, will or are likely to visit certain websites, and then infects these websites with malicious code waiting for the targets' trigger. The targets will be compromised e.g. when visit the websites, download software (malware) or click (malicious) links.	IDT, factors affecting trust and deception, social and organizational trust theory.	Fixed-action patterns, behavioral habits of site-visiting, think set and stereotyping, trust in familiar websites.
16	**Reverse social engineering**. The attacker sends an email using faked address to a new employee informing he / she that "a network test will be conduct recently, and if there is a network failure, please contact xxx". The attacker makes a network fault and waits for the new employee's request. After helping to resolve the problem, the attacker says sincerely "Would you like to do us a favor, just one minute, that completing a survey used for developing a security awareness training program for new employees; nearly 80% of the employees have already done this." "Ok, my pleasure." "Are you aware of our email policies? ... It can be dangerous to open unsolicited attachment ... We need to know your password to evaluate the security awareness of new employees. It is a secure matter" "Okay, it is ..."	Reciprocity norm, impression management theory, commitment and consistency, framing effect and cognitive bias, language invoke confusion - induce and manipulation, group influence and conformity, diffusion of responsibility, factors affecting trust and deception, IDT.	Inexperience, intuitive judgement, agreeableness, credulity, conformity, the desire to be helpful.

was made: we select the most complex attack scenario as an example and discuss it in great detail.

As a case in point, the reverse social engineering attack scenario (No. 16) is expounded as follows.

1) The attacker firstly sends an email using faked address (technical support department) to a new employee informing he / she that "a network test will be conduct recently, and if there is a network failure, please contact xxx-xxxx (the attacker's phone number)."
2) Then, the attacker makes a network fault and waits for the new employee's request.
3) Usually, new employees don't know many colleagues yet, and they don't know the procedures or the dos and don'ts of the organization (inexperience). When a network failure occurred, they call to the technical support using the number informed before.
4) After helping to resolve the problem, the attacker says sincerely "Would you like to do us a favor, just one minute, that completing a survey used for developing a security awareness and training program for new employees; nearly 80% of the employees have already done this."
5) In order to make a good first impression, new employees are eager to show how cooperative and quick to respond they can be (agreeableness, the desire to be helpful, conformity). This involves the impression management theory. With the influence of reciprocity norm, the attacker's help to resolve the problem portends the new employee's favor and commitment.
6) The benevolence of "security awareness and training program for new employees" and the sincere voice enhance the trust (intuitive judgement).
7) Low time cost ("just one minute") enhances the desire to be helpful. The group influence and cognitive bias of framing effect ("80% of the employees have already done this") lead to a conformity.
8) Thus, a commitment is obtained ("Ok, my pleasure").
9) The regular conversation that "Are you aware of our email policies? . . . It can be dangerous to open unsolicited attachment. . . " reflects the integrity and benevolence further. A high level of trust is likely obtained.
10) In this situation, "We need to know your password to evaluate the security awareness of new employees" maybe cause the new employee a slight worry, but "80% of the employees have already done this" lead to the diffusion of responsibility. Furthermore, the commitment and consistency compelling he / she continue the disclosure.
11) In addition, the expression that "It is a secure matter" not only means "know your password" is a matter about security (a routine that "to evaluate the security awareness of new employees"), but also implies that "know your password" is a secure matter without danger (which relieves the worry). This language expression evokes the thinking confusion state, in which the new employee's behavior and decision are induced and manipulated.
12) The attacker designs a great deal of strategic activities (interpersonal deception theory, IDT) and uses many factors affect trust and deception.
13) Ultimately, the new employee's password is compromised ("Okay, the password is . . . ").

VI. DISCUSSION

A. RELATED WORK

Social engineering is an interdisciplinary field which involves computer science, cybersecurity, psychology, social psychology, cognitive science, psycholinguistics, neuroscience, brain science, etc. In work [1], human vulnerabilities such as credulity, greed, ignorance, curiosity, carelessness, helpfulness have been mentioned. Yet only the human vulnerabilities are not sufficient to describe how social engineering attacks take effect. For effect mechanism, some works discussed or involved it in different context. Many scholars, e.g. [26], [78], [82]–[84], employ Cialdini's [5], [85] six principles of influence and persuasion (reciprocation, commitment and consistency, social proof, liking, authority, scarcity) to explain the success of social engineering attacks. Literature [86], [87] also discussed some psychological principles that exhibit some kind of power to influence or persuade people and take effect during a social engineering attack (strong affect, overloading, reciprocation, deceptive relationships, diffusion of responsibility and moral duty, authority, integrity and consistency). Mitnick and Simon [17] describes social engineering based on various kinds of deception. Stajano and Wilson [88] discussed seven principles of scam for system security (distraction, social compliance, herd, dishonesty, kindness, need and greed, time). Ferreira *et al.* [89] analyzed the relation (equal, include, overlap) among the above principles and presented a merged list of social engineering persuasion principles, i) authority, ii) social proof, iii) liking, similarity & deception, iv) commitment, reciprocation & consistency, v) distraction. However, the human vulnerabilities were not carefully concerned in these works, and other aspects of effect mechanisms are not involved.

B. ABOUT THE CONCEPTUAL MODEL

The conceptual model presented in Section II provides an integrative and structural perspective to understand how social engineering attacks work, rather than a single perspective. The model might be simple, yet it is also easy to understand. Although the model is not sufficient to constitute a domain ontology for social engineering, it identified three significant entities to get an insight into how social engineering attacks take effect. It conveys a concise idea that the attacker formulates certain attack scenarios to drive an organic combination of attack methods, effect mechanisms and human vulnerabilities, through which the attack process take effect to achieve the attack goal.

In addition, this model clarifies and avoids some mix-up among different entity types. For instance, impersonation, decoying, human vulnerabilities (friendliness, sympathy, ignorance) and six influence principles are treated as close-access techniques to exploit someone's trust in [51].

C. ABOUT THE LEVEL OF EFFECT MECHANISMS

Although some synthesized principles of persuasion were presented in [89], the underlying mechanisms were neglected. For instance, the second merged principle *social proof (sp)* consisted of three principles: *i) diffusion of responsibility and moral duty, ii) social proof and iii) herd*, and their logical relation was described as *i)* $\subset$ *iii)* $\subset$ *ii)*. However, 1) the underlying mechanism of diffusion of responsibility is that the group situation reduces the individual's evaluation apprehension, which offers the victims an excuse to avoid responsibility for their behaviors; 2) the underlying mechanism of principle social proof and herd is informational influence, in which the victims attempt to avoid unknown risks or seek the correct direction / behavior with the assumption that the actions (information) of group are correct; 3) moral duty is a kind of social norm in many cultures taking effect by normative influence: people are influenced to do something the norm requires due to the desire to be accepted or liked, regardless of their behavior is correct or not. Thus, a merged principle to ''constitute a basis for principles of social engineering'' in fact is based on three different underlying mechanisms.

We conducted an analysis of the effect mechanisms toward the fundamental level as much as possible, rather than a simply and upwards grouping. Hence, this paper offers a more clear explanation why the victims are exploited and why social engineering attacks become effective.

D. ABOUT THE COVERAGE AND COMPLETENESS

Besides the items mentioned in Section VI-A, this paper analyzed and discussed a wider range of effect mechanisms and human vulnerabilities. Overall, 30+ effect mechanisms in 6 aspects *(persuasion, social influence, cognition, attitude and behavior, trust and deception, language & thought and decision, emotion and decision)* and 40+ human vulnerabilities in 6 aspects *(cognition and knowledge, behavior and habit, emotion and feeling, human nature, personality traits, individual characteristics)* were summarized in Figure 2 (Appendix VII). Moreover, 16 attack scenarios together with these mechanisms and vulnerabilities are presented.

Nevertheless, did this paper provides a complete and exhaustive discussion of effect mechanisms, human vulnerabilities and attack methods for social engineering? The answer is 'No'. This is probably an unsolvable problem. Social engineering attacks not only exploit the obvious human vulnerabilities, but also the inconspicuous human factors. It seems every human factor involved provides the attacker a chance to turn it into a vulnerability. With the technology development and cyber-environment change, the attacker will create more attack scenarios, in which new attack methods are crafted, new effect mechanisms are found and more human vulnerabilities are exploited.

Even so, the presented mechanisms, vulnerabilities, scenarios and methods constitute plenty of materials for education, security awareness and training programs. Administrators, staffs, users and the public can use the proposed model as a knowledge schema of these materials. Both the material and model are helpful to increase the ability to understand and tackle with social engineering threat. And more attack scenarios can be generated based on the model and presented items. The education programs can be conducted by reminder, brochures, screensavers, courses, discussion, serious games, role-playing activities, penetration test, etc.

E. LIMITATION AND IMPLICATION

This paper analyzed and discussed many effect mechanisms and human vulnerabilities, 16 attack scenarios were also presented to illustrate their application. Although many of them are obvious effective or have been validated, there also some items are just theoretical feasible in the social engineering field (based on theoretical analysis and case study), i.e. they have not been empirical investigated. This is a limitation of this paper. Besides, the effectiveness of mechanisms and exploitability of human vulnerabilities may be affected by different environments, such as culture (individualism, collectivism), scenario (reality, cyberspace), medium (email, websites) and industry (IT or non-IT). And, empirical studies focusing on social engineering attacks is still relatively few. Thus, more empirical research is needed in the future. On the other hand, one of the merits of theoretical research might be it explorers a wider range and provides an integrative perspective. This paper offers lots of factors that can be further examined for future empirical research.

The conceptual model consists of 7 entities, but there are also some important entities have not been included, e.g. attack medium, and some relations among these entities have not been carefully defined. Besides, the relations among effect mechanisms, human vulnerabilities and attack methods are many-to-many, which might be clear displayed in the knowledge graph. Thus, in future work we will study the domain ontology of social engineering and its knowledge graph application.

VII. CONCLUSION

This paper proposes a conceptual model which provides an integrative and structural perspective to help the understanding of how social engineering attacks work. Three core entities (effect mechanisms, human vulnerabilities and attack methods) to get an insight into how social engineering attacks take effect are analyzed and discussed. A total of 30+ effect mechanisms and 40+ human vulnerabilities are summarized. Finally, 16 social engineering attack scenarios (which contains 13 attack methods) are presented to illustrate the application of these mechanisms, vulnerabilities and attack methods to understand how social engineering attacks work and take effect.

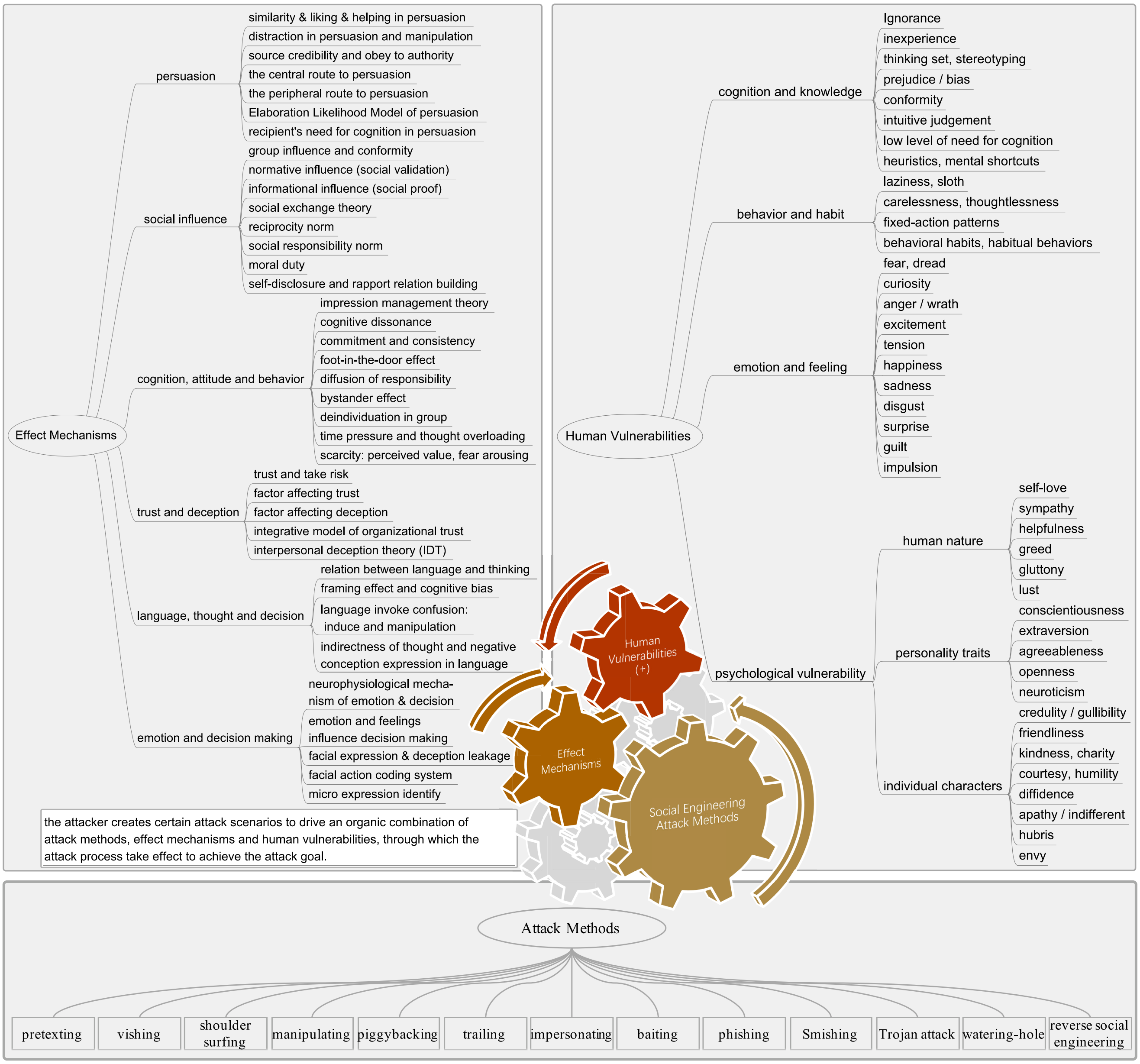
Effect Mechanisms
persuasion
similarity & liking & helping in persuasion
distraction in persuasion and manipulation
source credibility and obey to authority
the central route to persuasion
the peripheral route to persuasion
Elaboration Likelihood Model of persuasion
recipient's need for cognition in persuasion
social influence
group influence and conformity
normative influence (social validation)
informational influence (social proof)
social exchange theory
reciprocity norm
social responsibility norm
moral duty
self-disclosure and rapport relation building
cognition, attitude and behavior
impression management theory
cognitive dissonance
commitment and consistency
foot-in-the-door effect
diffusion of responsibility
bystander effect
deindividuation in group
time pressure and thought overloading
scarcity: perceived value, fear arousing
trust and deception
trust and take risk
factor affecting trust
factor affecting deception
integrative model of organizational trust
interpersonal deception theory (IDT)
language, thought and decision
relation between language and thinking
framing effect and cognitive bias
language invoke confusion: induce and manipulation
indirectness of thought and negative conception expression in language
emotion and decision making
neurophysiological mecha-nism of emotion & decision
emotion and feelings influence decision making
facial expression & deception leakage
facial action coding system
micro expression identify
the attacker creates certain attack scenarios to drive an organic combination of attack methods, effect mechanisms and human vulnerabilities, through which the attack process take effect to achieve the attack goal.
Human Vulnerabilities
cognition and knowledge
Ignorance
inexperience
thinking set, stereotyping
prejudice / bias
conformity
intuitive judgement
low level of need for cognition
heuristics, mental shortcuts
behavior and habit
laziness, sloth
carelessness, thoughtlessness
fixed-action patterns
behavioral habits, habitual behaviors
emotion and feeling
fear, dread
curiosity
anger / wrath
excitement
tension
happiness
sadness
disgust
surprise
guilt
impulsion
psychological vulnerability
human nature
self-love
sympathy
helpfulness
greed
gluttony
lust
personality traits
conscientiousness
extraversion
agreeableness
openness
neuroticism
individual characters
credulity / gullibility
friendliness
kindness, charity
courtesy, humility
diffidence
apathy / indifferent
hubris
envy
Human Vulnerabilities (+)
Effect Mechanisms
Social Engineering Attack Methods
Attack Methods
pretexting
vishing
shoulder surfing
manipulating
piggybacking
trailing
impersonating
baiting
phishing
Smishing
Trojan attack
watering-hole
reverse social engineering

FIGURE 2. Combined view of a schematic diagram of the proposed conceptual model and three mind maps of effect mechanisms, human vulnerabilities and attack methods.

APPENDIX

We create a combined view consisted of a schematic diagram of the proposed conceptual model and three mind maps of the effect mechanisms, human vulnerabilities and attack methods, to serve as a summary of the body of this paper. As Figure 2 shows.

REFERENCES

[1] Z. Wang, L. Sun, and H. Zhu, "Defining social engineering in cybersecurity," *IEEE Access*, vol. 8, pp. 85094–85115, 2020, doi: 10.1109/access.2020.2992807.

[2] D. Byrne, "An overview (and underview) of research and theory within the attraction paradigm," *J. Social Pers. Relationships*, vol. 14, no. 3, pp. 417–431, Jun. 1997.

[3] M. I. Norton, J. H. Frost, and D. Ariely, "Less is more: The lure of ambiguity, or why familiarity breeds contempt," *J. Personality Social Psychol.*, vol. 92, no. 1, p. 97, 2007.

[4] P. Miller, J. Kozu, and A. Davis, "Social influence, empathy, and prosocial behavior in cross-cultural perspective," in *Proc. Pract. Social Influence Multiple Cultures*, 2001, pp. 63–77.

[5] R. B. Cialdini, *Influence: Science and Practice*. Boston, MA, USA: Allyn and Bacon, 2001.

[6] P. R. Mims, J. J. Hartnett, and W. R. Nay, "Interpersonal attraction and help volunteering as a function of physical attractiveness," *J. Psychol.*, vol. 89, no. 1, pp. 125–131, Jan. 1975.

[7] S. G. West and T. Jan Brown, "Physical attractiveness, the severity of the emergency and helping: A field experiment and interpersonal simulation," *J. Exp. Social Psychol.*, vol. 11, no. 6, pp. 531–538, Nov. 1975.

[8] M. Harl. (1997). *People Hacking : The Psychology of Social Engineering*. [Online]. Available: http://www.textfiles.com/russian/cyberlib.narod.ru/lib/cin/se10.html

[9] P. C. Rosenblatt, ''Persuasion as a function of varying amounts of distraction,'' *Psychonomic Sci.*, vol. 5, no. 2, pp. 85–86, Feb. 1966, doi: 10.3758/BF03328292.

[10] R. A. Osterhouse and T. C. Brock, ''Distraction increases yielding to propaganda by inhibiting counterarguing,'' *J. Personality Social Psychol.*, vol. 15, no. 4, p. 344, 1970.

[11] R. E. Guadagno and R. B. Cialdini, *Online Persuasion and Compliance: Social Influence on the Internet and Beyond*. London, U.K.: Oxford Univ. Press, 2005, pp. 91–113.

[12] B. J. Sagarin, M. A. Britt, J. D. Heider, S. E. Wood, and J. E. Lynch, ''Bartering our attention: The distraction and persuasion effects of on-line advertisements,'' *Cognit. Technol.*, vol. 8, no. 2, pp. 4–17, 2003.

[13] D. R. Brandt, ''Listener propensity to counterargue, distraction, and resistance to persuasion,'' *Central States Speech J.*, vol. 30, no. 4, pp. 321–331, Dec. 1979.

[14] J. R. Priester and R. E. Petty, ''Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny,'' *Personality Social Psychol. Bull.*, vol. 21, no. 6, pp. 637–654, Jun. 1995, doi: 10.1177/0146167295216010.

[15] C. K. Hofling, E. Brotzman, S. Dalrymple, N. Graves, and C. M. Pierce, ''An experimental study in nurse-physician relationships,'' *The J. Nervous Mental Disease*, vol. 143, no. 2, pp. 171–180, 1966.

[16] S. Milgram and C. Gudehus, *Obedience to Authority*. New York, NY, USA: Ziff-Davis, 1978.

[17] K. D. Mitnick and W. L. Simon, *The Art of Deception: Controlling the Human Element of Security*. Hoboken, NJ, USA: Wiley, Aug. 2011.

[18] M. Butavicius, K. Parsons, M. Pattinson, and A. McCormac, ''Breaching the human firewall: Social engineering in phishing and spear-phishing emails,'' 2016, *arXiv:1606.00887*. [Online]. Available: http://arxiv.org/abs/1606.00887

[19] R. E. Petty, ''A cognitive response analysis of the temporal persistence of attitude changes induced by persuasive communications,'' Ph.D. dissertation, Graduate School, Ohio State Univ., Columbus, OH, USA, 1977. [Online]. Available: http://rave.ohiolink.edu/etdc/view?acc_num=osu1487955360601922

[20] R. E. Petty and J. T. Cacioppo, *Attitudes Persuasion: Classic Contemporary Approaches*. Dubuque, IA, USA: Wm. C. Brown, 1981.

[21] R. E. Petty and J. T. Cacioppo, *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. Berlin, Germany: Springer-Verlag, 1986.

[22] R. E. Petty and J. T. Cacioppo, ''The Elaboration Likelihood Model of Persuasion,'' in *Advances in Experimental Social Psychology*, vol. 19, L. Berkowitz, Ed. New York, NY, USA: Academic, Jan. 1986, pp. 123–205. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0065260108602142

[23] J. T. Cacioppo, R. E. Petty, J. A. Feinstein, and W. B. G. Jarvis, ''Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition,'' *Psychol. Bull.*, vol. 119, no. 2, p. 197, 1996.

[24] J. T. Cacioppo, R. E. Petty, C. F. Kao, and R. Rodriguez, ''Central and peripheral routes to persuasion: An individual difference perspective,'' *J. Personality Social Psychol.*, vol. 51, no. 5, p. 1032, 1986.

[25] T. R. Peltier, ''Social engineering: Concepts and solutions,'' *Inf. Syst. Secur.*, vol. 15, no. 5, pp. 13–21, Nov. 2006.

[26] M. Workman, ''Gaining access with social engineering: An empirical study of the threat,'' *Inf. Syst. Secur.*, vol. 16, no. 6, pp. 315–331, Dec. 2007.

[27] D. G. Myers, *Social Psychology*, 11th ed. New York, NY, USA: McGraw-Hill, 2012.

[28] S. Milgram, L. Bickman, and L. Berkowitz, ''Note on the drawing power of crowds of different size.,'' *J. Personality Social Psychol.*, vol. 13, no. 2, pp. 79–82, 1969.

[29] M. Deutsch and H. B. Gerard, ''A study of normative and informational social influences upon individual judgment,'' *The J. Abnormal Social Psychol.*, vol. 51, no. 3, p. 629, 1955.

[30] U. G. Foa and E. B. Foa, *Resource Theory Social Exchange*. New York, NY, USA: General Learning Press, 1975.

[31] A. W. Gouldner, ''The norm of reciprocity: A preliminary statement,'' *Amer. Sociol. Rev.*, vol. 4, pp. 161–178, Dec. 1960.

[32] L. Berkowitz, ''Social norms, feelings, and other factors affecting helping and altruism,'' in *Advances in experimental social psychology*, vol. 6. Amsterdam, The Netherlands: Elsevier, 1972, pp. 63–108.

[33] J. Baron and J. G. Miller, ''Limiting the scope of moral obligations to help: A cross-cultural investigation,'' *J. Cross-Cultural Psychol.*, vol. 31, no. 6, pp. 703–725, Nov. 2000.

[34] V. J. Derlaga and J. H. Berg, *Self-Disclosure: Theory, Research, and Therapy*. Cham, Switzerland: Springer, 1987.

[35] N. L. Collins and L. C. Miller, ''Self-disclosure and liking: A meta-analytic review,'' *Psychol. Bull.*, vol. 116, no. 3, p. 457, 1994.

[36] L. C. Miller, J. H. Berg, and R. L. Archer, ''Openers: Individuals who elicit intimate self-disclosure,'' *J. Personality Social Psychol.*, vol. 44, no. 6, p. 1234, 1983.

[37] S. M. Jourard, *The Transparent Self: Self-Disclosure and Well-Being*, vol. 17. Princeton, NJ, USA: Van Nostrand, 1964.

[38] M. R. Leary, *Self-Presentation: Impression Management and Interpersonal Behavior*. Evanston, IL, USA: Routledge, 2019.

[39] L. Festinger, *A Theory Cognition Dissonance*. Stanford, CA, USA: Stanford Univ. Press, 1962, vol. 2.

[40] J. L. Freedman and S. C. Fraser, ''Compliance without pressure: The foot-in-the-door technique,'' *J. Personality Social Psychol.*, vol. 4, no. 2, p. 195, 1966.

[41] B. Latanã and J. M. Dabbs, Jr., ''Sex, group size and helping in three cities,'' *Sociometry*, vol. 1, pp. 180–194, Jun. 1975.

[42] B. Latanã and S. Nida, ''Ten years of research on group size and helping,'' *Psychol. Bull.*, vol. 89, no. 2, p. 308, 1981.

[43] J. M. Darley and B. Latana, ''Bystander intervention in emergencies: Diffusion of responsibility.'' *J. Personality Social Psychol.*, vol. 8, no. 4p1, p. 377, 1968.

[44] M. Levine and S. Crowther, ''The responsive bystander: How social group membership and group size can encourage as well as inhibit bystander intervention,'' *J. Personality Social Psychol.*, vol. 95, no. 6, p. 1429, 2008.

[45] C. A. Blair, L. Foster Thompson, and K. L. Wuensch, ''Electronic helping behavior: The virtual presence of others makes a difference,'' *Basic Appl. Social Psychol.*, vol. 27, no. 2, pp. 171–178, Jun. 2005.

[46] L. Festinger, A. Pepitone, and T. Newcomb, ''Some consequences of de-individuation in a group,'' *The J. Abnormal Social Psychol.*, vol. 47, no. 2, pp. 382–389, 1952.

[47] S. G. Harkins and J. M. Jackson, ''The role of evaluation in eliminating social loafing,'' *Personality Social Psychol. Bull.*, vol. 11, no. 4, pp. 457–465, Dec. 1985.

[48] N. J. Evans, ''Information technology social engineering: An academic definition and study of social engineering-analyzing the human firewall,'' Ph.D. dissertation, Elect. Comput. Eng., Iowa State Univ., Ames, IA, USA, 2009.

[49] S. M. Albladi and G. R. S. Weir, ''Predicting individuals' vulnerability to social engineering in social networks,'' *Cybersecurity*, vol. 3, no. 1, pp. 1–19, Dec. 2020.

[50] A. Chitrey, D. Singh, and V. Singh, ''A comprehensive study of social engineering based attacks in india to develop a conceptual model,'' *Int. J. Inf. Netw. Secur. (IJINS)*, vol. 1, no. 2, p. 45, Jun. 2012.

[51] L. Laribee, ''Development of methodical social engineering taxonomy project,'' M.S. thesis, Dept. Comput. Sci., Naval Postgraduate School, Monterey, CA, USA, Jun. 2006. [Online]. Available: https://calhoun.nps.edu/handle/10945/2734

[52] R. C. Mayer, J. H. Davis, and F. D. Schoorman, ''An integrative model of organizational trust,'' *Acad. Manage. Rev.*, vol. 20, no. 3, pp. 709–734, Jul. 1995.

[53] P. Sztompka, *Trust: A Sociological Theory*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[54] J. K. Burgoon and D. B. Buller, *Interpersonal Deception Theory*. Atlanta, GA, USA: American Cancer Society, Dec. 2015, pp. 1–6.

[55] P. W. Linville, G. W. Fischer, and B. Fischhoff, *AIDS Risk Perceptions and Decision Biases*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, 1993, pp. 5–38.

[56] Paranoiahax. (Mar. 2009). *Comment of NLP and Social Engineering—Hacking the Human Mind Article*.[Online]. Available: https://www.hellboundhackers.org/articles/read-article.php?article_id=8%78

[57] J. E. LeDoux and R. Brown, ''A higher-order theory of emotional consciousness,'' *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 10, pp. E2016–E2025, Mar. 2017.

[58] E. A. Phelps, K. M. Lempert, and P. Sokol-Hessner, ''Emotion and decision making: Multiple modulatory neural circuits,'' *Annu. Rev. Neurosci.*, vol. 37, no. 1, pp. 263–287, Jul. 2014.

[59] R. Raghunathan and M. T. Pham, ''All negative moods are not equal: Motivational influences of anxiety and sadness on decision making,'' *Org. Behav. Hum. Decis. Processes*, vol. 79, no. 1, pp. 56–77, Jul. 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0749597899928388

[60] J. S. Lerner and D. Keltner, ''Fear, anger, and risk,'' *J. Personality Social Psychol.*, vol. 81, no. 1, pp. 146–159, 2001.

[61] A. M. Isen and B. Means, "The influence of positive affect on decision-making strategy," *Social Cognition*, vol. 2, no. 1, pp. 18–31, Mar. 1983.

[62] R. E. Petty, D. W. Schumann, S. A. Richman, and A. J. Strathman, "Positive mood and persuasion: Different roles for affect under high- and low-elaboration conditions.," *J. Personality Social Psychol.*, vol. 64, no. 1, pp. 5–20, Jan. 1993.

[63] H. Y. Jin, "The causes and laws of unconsciousness conditioned emotional response effect: 187 case studies," *Psychol. Explor.*, vol. 4, pp. 36–39, May 1992.

[64] D. B. Buller and J. K. Burgoon, "Interpersonal deception theory," *Commun. Theory*, vol. 6, no. 3, pp. 203–242, Aug. 1996.

[65] P. Ekman, W. V. Friesen, and J. C. Hager, *Facial action coding system: The manual on CD-ROM*, 2nd ed. Salt Lake City, Utah: A Human Face, 2002.

[66] P. Ekman, *Micro Expression Training Tool (METT) and Subtle Expression Training Tool (SETT)*. San Francisco, CA, USA: Paul Ekman Company, 2003.

[67] P. Schaab, K. Beckers, and S. Pape, "A systematic gap analysis of social engineering defence mechanisms considering social psychology," in *Proc. HAISA*, Jul. 2016, pp. 241–251.

[68] E. D. Frangopoulos, M. M. Eloff, and L. M. Venter, "Psychosocial risks: Can their effects on the security of information systems really be ignored?" in *Proc. HAISA*, 2012, pp. 52–63.

[69] R. Gulati, "The threat of social engineering and your defense against it," SANS Inst., Bethesda, MD, USA, Tech. Rep. 1232, 2003. [Online]. Available: https://www.sans.org/reading-room/whitepapers/engineering/paper/1232

[70] M. Nohlberg, "Social engineering: Understanding, measuring and protecting against attacks," Univ. Skãvde, Sweden, Stockholm, Tech. Rep., 2007.

[71] E. M. Maseno, "Vishing attack detection model for mobile users," M.S. thesis, Comput. Inf. Manage., KCA Univ., Nairobi, Kenya, Nov. 2017. [Online]. Available: http://41.89.49.13:8080/xmlui/handle/123456789/1276

[72] M. Junger, L. Montoya, and F.-J. Overink, "Priming and warnings are not effective to prevent social engineering attacks," *Comput. Hum. Behav.*, vol. 66, pp. 75–87, Jan. 2017.

[73] D. Harley, "Re-floating the titanic: Dealing with social engineering attacks," in *Proc. EICAR*, 1998, pp. 4–29.

[74] A. Thapar, "Social engineering: An attack vector most intricate to tackle," CISSP, Clearwater, FL, USA, Tech. Rep. 106841, 2007. [Online]. Available: http://www.infosecwriters.com/text_resources/pdf/Social_Engineering_ATh%apar.pdf

[75] M. Bezuidenhout, F. Mouton, and H. S. Venter, "Social engineering attack detection model: SEADM," in *Proc. Inf. Secur. South Afr.*, Aug. 2010, pp. 1–8.

[76] D. Reading, "Social engineering, the USB way," Dark Reading, New York, NY, USA, Tech. Rep. 1128081, 2006. [Online]. Available: https://www.darkreading.com/attacks-breaches/social-engineering-the-usb%-way/d/d-id/1128081

[77] W. Fan, K. Lwakatare, and R. Rong, "Social engineering: IE based model of human weakness to investigate attack and defense," *SCIREA J. Inf. Sci. Syst. Sci.*, vol. 1, no. 2, pp. 34–57, 2016.

[78] S. Uebelacker and S. Quiel, "The social engineering personality framework," in *Proc. Workshop Socio-Tech. Aspects Secur. Trust*, Jul. 2014, pp. 24–30.

[79] J. Stewart and M. Dawson, "How the modification of personality traits leave one vulnerable to manipulation in social engineering," *Int. J. Inf. Privacy, Secur. Integrity*, vol. 3, no. 3, pp. 187–208, Jan. 2018.

[80] N. Tsinganos, G. Sakellariou, P. Fouliras, and I. Mavridis, "Towards an automated recognition system for chat-based social engineering attacks in enterprise environments," in *Proc. 13th Int. Conf. Availability, Rel. Secur.*, New York, NY, USA, Aug. 2018, p. 53, doi: 10.1145/3230833.3233277.

[81] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *J. Personality*, vol. 60, no. 2, pp. 175–215, Jun. 1992.

[82] A. Algarni, Y. Xu, T. Chan, and Y.-C. Tian, "Toward understanding social engineering," in *Proc. 8th Int. Conf. Legal, Secur. Privacy*. Bangkok, Thailand: The International Association of IT Lawyers (IAITL), Nov. 2013, pp. 279–300. [Online]. Available: https://eprints.qut.edu.au/67479/

[83] F. Mouton, L. Leenen, and H. S. Venter, "Social engineering attack examples, templates and scenarios," *Comput. Secur.*, vol. 59, no. 2, pp. 186–209, Jun. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167404816300268

[84] J. V. Rensburg and K. Shandre, "The human element in information security : An analysis of social engineering attacks in the greater Tshwane area of Gauteng, South Africa," M.S. thesis, Dept. Criminol. Secur. Sci., Univ. South Africa, Grahamstown, South Africa, Jun. 2017. [Online]. Available: http://uir.unisa.ac.za/handle/10500/22681

[85] R. Cialdini, *Influence: The Psychology of Persuasion* (Collins Business Essentials). New York, NY, USA: Harper Collins, 2009. [Online]. Available: https://books.google.com.br/books?id=5dfv0HJ1TEoC

[86] D. Gragg, "A multi-level defense against social engineering," *SANS Reading Room*, vol. 13, p. 15, Mar. 2003. [Online]. Available: http://taupe.free.fr/book/psycho/social%20engineering/Social%20Engineer%ing%20-%20Sans%20Institute%20-%20Multi%20Level%20Defense%20Against%20Social%20%Engineering.pdf

[87] B. Oosterloo, "Managing social engineering risk: Making social engineering transparant," Ph.D. dissertation, Ind. Eng. Manage., University of Twente, Enschede, The Netherlands, 2008. [Online]. Available: http://essay.utwente.nl/59233/1/scriptie_B_Oosterloo.pdf

[88] F. Stajano and P. Wilson, "Understanding scam victims: Seven principles for systems security," *Commun. ACM*, vol. 54, no. 3, pp. 70–75, Mar. 2011, doi: 10.1145/1897852.1897872.

[89] A. Ferreira, L. Coventry, and G. Lenzini, "Principles of persuasion in social engineering and their use in phishing," in *Proc. Int. Conf. Hum. Aspects Inf. Secur., Privacy, Trust*. Cham, Switzerland: Springer, 2015, pp. 36–47.

ZUOGUANG WANG received the master's degree from Information Engineering University. He is currently pursuing the Ph.D. degree with the University of Chinese Academy of Sciences (UCAS). He is also pursuing the Ph.D. degree in cybersecurity with the Institute of Information Engineering, Chinese Academy of Sciences. He has participated in one of the National Key Research and Development Programs of China. His main research interests include defense against social engineering, industrial control system security, and the Internet-of-Things (IoT) security.

HONGSONG ZHU (Member, IEEE) received the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences. He is currently a Professor with the Institute of Information Engineering, Chinese Academy of Sciences, and the University of Chinese Academy of Sciences. His main research interests include network measurement, the Internet-of-Things (IoT) security, and defense against social engineering. He is also a Senior Member of the China Computer Federation and a member of the Select Committee of China Computer Federation Technical Commission on Sensor Network (CWSN).

LIMIN SUN is currently a Professor with the Institute of Information Engineering, Chinese Academy of Sciences, and the University of Chinese Academy of Sciences. His main research interests include the Internet-of-Things (IoT) security, industrial control system security, and defense against social engineering. He is an Editor of the *Journal of Computer Science* and the *Journal of Computer Applications*, and a Guest Editor of special issues in *EURASIP* and the *Journal of Networks*. He is also the Secretary General of the Select Committee of CWSN.